International Journal of Medicine
pISSN 1468-3814 | eISSN 2667-7008
https://theinternationalmedicine.com/

Original Researcher Article *DOI: 10.61336/im/26-08-14*

Diagnostic Accuracy of the Bedside Index for Severity in Acute Pancreatitis (BISAP) Against the Modified Computed Tomography Severity Index: A Prospective Observational Study.

Dr. SOWMYA G R[1]*, Dr. Prathvi Nandalike[2], Dr. Manjunath[3].
[1]ASSISTANT PROFESSOR DEPT OF GENERAL MEDICINE BMCRC, BALLARI.
[2]Assistant professor JGMMMC KLE HUBBALLI.
[3]Assistant professor Dept of GENERAL MEDICINE KIMS Koppala.

Correspondence:
Dr. Sowmya G R.
ASSISTANT PROFESSOR DEPT OF GENERAL MEDICINE BMCRC, BALLARI.

Email:
sowmyagr27@gmail.com

Received: 01-07-2026
Revised: 11-07-2026
Accepted: 20-07-2026
Published: 03-08-2026

Abstract

Background: Early identification of patients with severe acute pancreatitis (AP) permits timely intensive-care admission, aggressive fluid resuscitation and close monitoring for organ failure. Contrast-enhanced computed tomography (CECT) with the modified CT severity index (mCTSI) remains the radiological reference standard but is expensive, requires iodinated contrast and is frequently unavailable in resource-limited settings. The Bedside Index for Severity in Acute Pancreatitis (BISAP) is a five-point clinical score computable within 24 hours of admission. ***Objectives:*** To determine the sensitivity, specificity, predictive values and discriminative performance of the BISAP score against the mCTSI for the identification of severe AP, and to identify the optimal BISAP threshold in this population. ***Methods:*** A hospital-based prospective observational study enrolled 50 consecutive adults with AP admitted to the Department of General Medicine, Ballari Medical College and Research Centre, Ballari, between June 2022 and December 2023. BISAP was calculated within 24 hours of admission; CECT of the abdomen and pelvis was performed on days 3–7 and scored using the mCTSI by a radiologist blinded to the BISAP score. An mCTSI of 8–10 defined radiologically severe disease. Diagnostic indices were computed with Wilson 95% confidence intervals (CI); discrimination was assessed by the area under the receiver operating characteristic curve (AUC). ***Results:*** Mean age was 33.82 ± 8.03 years and 49 of 50 patients (98.0%) were male. Nine patients (18.0%) had mCTSI-severe disease. At the conventional threshold of BISAP ≥ 3, sensitivity was 44.4% (95% CI 18.9–73.3), specificity 97.6% (95% CI 87.4–99.6), positive predictive value 80.0% (95% CI 37.6–96.4), negative predictive value 88.9% (95% CI 76.5–95.2) and overall accuracy 88.0% (95% CI 76.2–94.4); Fisher's exact p = 0.003. The AUC of BISAP as an ordinal predictor was 0.799 (95% CI 0.616–0.983, p = 0.004). Lowering the threshold to BISAP ≥ 2 raised sensitivity to 77.8% with specificity 61.0% and a negative predictive value of 92.6%. BISAP and mCTSI were significantly correlated (Pearson r = 0.657, p < 0.001). ***Conclusions:*** In this cohort BISAP was a highly specific but insensitive rule-in test for radiologically severe AP at the conventional cut-off of ≥ 3. Its overall discrimination was good (AUC 0.80), and a lower threshold of ≥ 2 offered a better rule-out profile. BISAP should be regarded as a triage adjunct that identifies high-risk patients early, not as a substitute for CECT.

Keywords: Acute pancreatitis; BISAP; Modified CT severity index; Diagnostic accuracy; Severity stratification.

INTRODUCTION

Acute pancreatitis (AP) is an acute inflammatory disorder of the pancreas initiated by the premature intra-acinar activation of trypsinogen, resulting in autodigestion of the gland and a variable systemic inflammatory response.¹ The clinical spectrum is broad, ranging from self-limiting interstitial oedematous disease to necrotising pancreatitis complicated by persistent multi-organ failure. Overall mortality is approximately 1%, but rises to 20–30% among patients with severe acute pancreatitis (SAP).² Population-level incidence data from India are sparse; reported incidence from Western populations ranges from 41.9 to 49.3 per 100,000 population in Scotland, Finland and the United States respectively, and tertiary Indian centres report of the order of 55 new cases annually.³

The revised Atlanta classification of 2012 requires two of three criteria for diagnosis — characteristic epigastric pain, serum amylase or lipase at least three times the upper limit of normal, and cross-sectional imaging consistent with AP — and stratifies severity into mild (no organ failure, no complications), moderately severe (transient organ failure of less than 48 hours and/or local complications) and severe (persistent organ failure exceeding 48 hours).^{4,5} Because no

disease-modifying pharmacotherapy exists, management rests on fluid resuscitation, analgesia, early enteral nutrition and organ support, all of which are most effective when instituted before organ failure becomes established.^{6,7} Severity prediction is therefore not an academic exercise but the principal determinant of the level of care a patient receives. Numerous prognostic instruments have been developed. Ranson's criteria, described in 1977, classify severity reasonably well but require 48 hours to complete, closing the early therapeutic window.⁸

The Acute Physiology and Chronic Health Evaluation (APACHE) II score is available within 24 hours but is cumbersome and requires numerous laboratory variables.⁹ Radiologically, Balthazar and colleagues introduced the CT severity index (CTSI) in 1990 by combining a morphological grade with the extent of pancreatic necrosis,¹⁰ and Mortele and colleagues subsequently proposed a simplified modified CT severity index (mCTSI) in 2004 that incorporates extrapancreatic complications and correlates more closely with length of stay, need for intervention, organ failure and death.¹¹ Contrast-enhanced CT nevertheless requires intravenous iodinated contrast, carries a radiation burden, is optimally performed only after 72 hours, and is often unaffordable or unavailable in district and rural hospitals.

Against this background, Wu and colleagues derived the Bedside Index for Severity in Acute Pancreatitis (BISAP) in 2008 from a large population-based cohort.¹² BISAP assigns one point each for blood urea nitrogen above 25 mg/dL, impaired mental status, systemic inflammatory response syndrome, age over 60 years and pleural effusion on imaging, all ascertainable within 24 hours of presentation.¹³ Validation studies have reported widely divergent operating characteristics: Khanna et al. found a sensitivity of 74.2% and specificity of 68.3%, whereas Chen et al. reported 61.4% and 83.1% respectively.^{14,15} Comparative studies against radiological indices have likewise reached opposing conclusions, some favouring CTSI¹⁶ and others favouring BISAP.¹⁷ Such heterogeneity suggests that BISAP performance is population-dependent and requires local validation. We therefore prospectively compared BISAP against the mCTSI in a south Indian cohort, with the specific aims of quantifying its diagnostic indices with confidence intervals, estimating its discriminative ability, and determining the threshold most useful for early triage.

MATERIALS AND METHODS

This was a hospital-based prospective observational study conducted in the Department of General Medicine, Ballari Medical College and Research Centre (Vijayanagar Institute of Medical Sciences), Ballari, Karnataka, India, a government tertiary-care teaching hospital serving a predominantly rural catchment population in north Karnataka. Recruitment ran from June 2022 to December 2023.

Participants. Consecutive patients presenting to the casualty department and admitted with a diagnosis of AP were screened. Diagnosis required at least two of the following three criteria in accordance with the revised Atlanta classification: characteristic epigastric abdominal pain with or without radiation to the back; serum amylase or serum lipase elevated to at least three times the upper limit of normal; and imaging findings consistent with AP.⁴ Patients aged above 15 years were eligible.

Exclusion criteria were: age 15 years or below; pre-existing chronic kidney disease, since an elevated blood urea nitrogen attributable to renal rather than pancreatic pathology would spuriously inflate the BISAP score; acute-on-chronic pancreatitis; and presentation more than 24 hours after symptom onset, which would preclude valid computation of the 24-hour BISAP score.

Sample size. The required sample size was estimated using the formula $n = Z^2\alpha \bullet P(1 - P) / d^2$, with $Z\alpha = 1.96$ for a two-sided 95% confidence level, P the expected proportion in the population and d the absolute permissible error. Assuming an expected proportion of 4.1% derived from a previous Indian study and accepting a 6% absolute margin of error, the minimum requirement was 48 cases. Fifty patients were recruited to allow for incomplete data.

Index test. The BISAP score was computed for every patient within 24 hours of admission by the treating physician. One point was assigned for each of: blood urea nitrogen greater than 25 mg/dL; impaired mental status (Glasgow Coma Scale below 15); systemic inflammatory response syndrome, defined as two or more of temperature below 36 °C or above 38 °C, respiratory rate above 20 breaths/min or $PaCO_2$ below 32 mmHg, pulse above 90 beats/min, and white cell count below 4,000 or above 12,000 cells/mm^3 or more than 10% immature bands; age above 60 years; and pleural effusion on imaging. The composite score ranged from 0 to 5.^{12,13} Complete blood count with peripheral smear, renal function tests and chest radiography were performed on admission to populate these variables.

Reference standard. Contrast-enhanced CT of the abdomen and pelvis was performed between days 3 and 7 of admission and scored using the modified CT severity index of Mortele et al.¹¹ Pancreatic inflammation scored 0

(normal), 2 (intrinsic abnormality with or without peripancreatic fat inflammatory change) or 4 (pancreatic or peripancreatic fluid collection or fat necrosis); pancreatic necrosis scored 0 (none), 2 (≤ 30%) or 4 (> 30%); and the presence of any extrapancreatic complication (pleural effusion, ascites, vascular complications, parenchymal complications or gastrointestinal involvement) scored 2. Total scores of 0–2, 4–6 and 8–10 denoted mild, moderate and severe disease respectively. Scans were reported by consultant radiologists who were blinded to the BISAP score.

Statistical analysis. Data were analysed using IBM SPSS Statistics for Windows version 29.0. Categorical variables were summarised as frequencies and percentages, continuous variables as mean ± standard deviation. An mCTSI of 8–10 was taken as the binary reference standard for severe disease. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and accuracy were calculated from the 2 × 2 contingency table at BISAP thresholds of ≥ 1, ≥ 2 and ≥ 3, with Wilson score 95% confidence intervals.

Association was tested by Fisher's exact test. Discrimination was quantified by the area under the receiver operating characteristic curve, computed as the Mann–Whitney statistic for BISAP treated as an ordinal predictor, with Hanley–McNeil standard errors. The Youden index (sensitivity + specificity − 1) identified the optimal threshold. Correlation between BISAP and mCTSI was assessed by Pearson's coefficient. A two-sided p value below 0.05 was considered statistically significant.

Ethics. The study was approved by the Institutional Ethics Committee, Vijayanagar Institute of Medical Sciences, Ballari (Approval No. 50/PG Thesis/2021-22, dated 27 January 2023). Written informed consent was obtained from every participant or an authorised representative.

RESULTS

Baseline characteristics

Fifty patients fulfilled the eligibility criteria and completed both index test and reference standard; there were no losses to follow-up and no indeterminate CT results.

Table 1. Baseline demographic and clinical characteristics (n = 50)

Characteristic	Value
Age, years — mean ± SD	33.82 ± 8.03
Age range, years	16–50
Male — n (%)	49 (98.0)
Female — n (%)	1 (2.0)
BISAP score — mean ± SD	1.38 ± 0.97
mCTSI — mean ± SD	5.60 ± 1.85
mCTSI-severe disease (score 8–10) — n (%)	9 (18.0)
BISAP ≥ 3 — n (%)	5 (10.0)

The cohort was strikingly young and male-predominant, with a mean age well below the fifth-to-sixth-decade peak described in Western series. No patient in the cohort was older than 60 years, meaning that the age component of the BISAP score contributed zero points throughout and the effective score range was 0–4.

Distribution of index test and reference standard

Table 2. Distribution of BISAP and modified CT severity index scores (n = 50)

BISAP score	n (%)		mCTSI	n (%)	Mortele category
0	10 (20.0)		2	2 (4.0)	Mild
1	17 (34.0)		4	18 (36.0)	Moderate
2	18 (36.0)		6	21 (42.0)	Moderate
3	4 (8.0)		8	6 (12.0)	Severe
4	1 (2.0)		10	3 (6.0)	Severe
5	0 (0)				

Only two patients (4.0%) had radiologically mild disease, 39 (78.0%) moderate and 9 (18.0%) severe. The concentration of the cohort in the moderate band reflects the requirement for a CECT-confirmed diagnosis and the referral pattern of a tertiary centre.

Diagnostic performance at the conventional threshold

Table 3. Cross-tabulation of BISAP ≥ 3 against mCTSI-severe disease

	mCTSI severe (8–10)	mCTSI not severe (0–6)	Total
BISAP ≥ 3	4 (TP)	1 (FP)	5
BISAP < 3	5 (FN)	40 (TN)	45
Total	9	41	50

Fisher's exact test p = 0.003.

Table 4. Diagnostic indices of BISAP ≥ 3 for mCTSI-severe acute pancreatitis

Index	Estimate	95% CI (Wilson)
Sensitivity	44.4%	18.9–73.3
Specificity	97.6%	87.4–99.6
Positive predictive value	80.0%	37.6–96.4
Negative predictive value	88.9%	76.5–95.2
Overall accuracy	88.0%	76.2–94.4
Positive likelihood ratio	18.2	—
Negative likelihood ratio	0.57	—
Youden index	0.420	—

The dominant feature of the table is asymmetry. A BISAP score of 3 or more was rarely wrong when positive — only one of five such patients had a non-severe CT — giving a positive likelihood ratio of 18.2 and making it a powerful rule-in signal. Conversely, five of the nine radiologically severe patients scored below 3 and would have been missed by this threshold alone. The high accuracy of 88.0% is driven largely by the low prevalence of severe disease (18.0%) and should not be read as evidence of a well-balanced test.

Threshold analysis

Table 5. Performance of BISAP across candidate thresholds

Threshold	TP	FP	FN	TN	Sensitivity	Specificity	PPV	NPV	Accuracy	Youden	Fisher's p
BISAP ≥ 1	9	31	0	10	100.0%	24.4%	22.5%	100.0%	38.0%	0.244	0.174
BISAP ≥ 2	7	16	2	25	77.8%	61.0%	30.4%	92.6%	64.0%	0.388	0.062
BISAP ≥ 3	4	1	5	40	44.4%	97.6%	80.0%	88.9%	88.0%	0.420	0.003

The Youden index was maximised at BISAP ≥ 3 (0.420), narrowly exceeding BISAP ≥ 2 (0.388). However, the two thresholds serve different clinical purposes. A score of 0 excluded severe disease entirely in this cohort (NPV 100%), and a threshold of ≥ 2 retained a negative predictive value of 92.6% while capturing more than three-quarters of severe cases — an operating point better suited to a screening decision about whether to arrange CECT or arrange transfer.

Discrimination and correlation

Treated as an ordinal predictor rather than dichotomised, BISAP discriminated well between radiologically severe and non-severe disease, with an area under the receiver operating characteristic curve of 0.799 (95% CI 0.616–0.983; p = 0.004). BISAP and mCTSI were significantly and positively correlated (Pearson r = 0.657, p < 0.001; Spearman ρ = 0.692, p < 0.001). Patients with severe disease were also significantly older than those without (39.0 ± 10.2 versus 32.7 ± 7.1 years; p = 0.031), despite no patient reaching the 60-year BISAP age threshold.

DISCUSSION

This prospective study evaluated the BISAP score against the modified CT severity index in 50 patients with acute pancreatitis and found a distinctly asymmetric performance profile: at the conventional threshold of 3 or more, BISAP was highly specific (97.6%) but poorly sensitive (44.4%), with a positive likelihood ratio of 18.2 and a negative likelihood ratio of 0.57. Treated as an ordinal variable, however, the score discriminated well (AUC 0.799) and correlated significantly with mCTSI (r = 0.657, p < 0.001). The practical implication is that a high BISAP score should prompt immediate escalation, whereas a low score cannot safely be used in isolation to withhold further assessment.

Our AUC of 0.80 sits comfortably within the published range. Khanna et al. reported an AUC of 0.80 for BISAP in a comparative evaluation of eight scoring systems,¹⁴ Chen et al. 0.762 in a Chinese cohort,¹⁵ and Chatterjee et al. 0.811 for BISAP ≥ 2 in an Indian intensive-care population, concluding that BISAP was comparable to APACHE II and mCTSI in accuracy while being substantially simpler to compute.¹⁸ Papachristou et al., comparing BISAP with Ranson's, APACHE II and CTSI, found broadly equivalent performance across all four and argued

that simple multifactorial scores may have reached the ceiling of their predictive utility.¹⁹ Cho et al. reached the same conclusion.²⁰ Our data are consistent with this literature: BISAP is useful but not sufficient.

The low sensitivity we observed deserves particular scrutiny, because it is largely explicable by the demographics of our cohort. No patient exceeded 60 years of age, so one of the five BISAP components was structurally uninformative and the attainable score range was compressed to 0–4. Studies drawn from older populations, in which the age criterion contributes meaningfully, would be expected to report higher sensitivity at the same threshold. Gupta et al., whose cohort had a mean age of 46.5 years, reported a sensitivity of 90.9% and specificity of 85.7%,¹⁶ and Kim et al. reported 79.2% and 88.5% respectively. This ceiling effect is an important and underappreciated limitation of applying BISAP unmodified to the young, alcohol-associated pancreatitis populations common in parts of India, and argues for local threshold calibration rather than uncritical adoption of the derivation-cohort cut-off.

A second consideration is that BISAP and mCTSI measure related but non-identical constructs. BISAP is a clinical and physiological score capturing the systemic inflammatory response, whereas mCTSI quantifies pancreatic and peripancreatic morphology. A patient may have extensive necrosis without immediate systemic derangement, or profound systemic inflammation with modest radiological change. Discordance between the two is therefore expected rather than anomalous, and the correlation coefficient of 0.657 — accounting for roughly 43% of shared variance — is best interpreted as confirming that the two indices track a common underlying severity while retaining independent information. Tahir et al. made a similar argument in showing that the neutrophil-to-lymphocyte ratio agreed more closely with the revised Atlanta classification than mCTSI did.²¹

The principal limitations are the small sample size, which produces wide confidence intervals — the sensitivity estimate spans 18.9% to 73.3% and cannot exclude either poor or acceptable performance — the single-centre design, the marked sex imbalance (98% male) that limits generalisability to women, and the absence of hard clinical endpoints such as organ failure, intensive-care admission, length of stay or mortality. Because CECT was used as the reference standard rather than the revised Atlanta clinical classification, our findings speak to radiological rather than clinical severity. We also did not compare BISAP against APACHE II or Ranson's score, so relative performance among clinical indices remains untested here.

CONCLUSION

In this prospective cohort of 50 predominantly young male patients with acute pancreatitis, the BISAP score demonstrated good overall discrimination against the modified CT severity index (AUC 0.799) and a significant positive correlation with it ($r = 0.657$, $p < 0.001$). At the conventional threshold of 3 or more it functioned as a highly specific rule-in test (specificity 97.6%, positive likelihood ratio 18.2) but lacked sensitivity (44.4%), in part because no patient met the 60-year age criterion. A lower threshold of 2 or more offered a more favourable rule-out profile (sensitivity 77.8%, negative predictive value 92.6%) and may be the more appropriate trigger for arranging cross-sectional imaging or interfacility transfer. BISAP is best positioned as an early triage adjunct in settings where CECT is delayed or unavailable, and should complement rather than replace radiological and clinical severity assessment. Larger multicentre studies with hard clinical endpoints and locally calibrated thresholds are required before BISAP can be recommended as a standalone decision instrument.

REFERENCES

1. Szatmary P, Grammatikopoulos T, Cai W, Huang W, Mukherjee R, Halloran C, et al. Acute pancreatitis: diagnosis and treatment. Drugs. 2022;82(12):1251–76.
2. Werge M, Novovic S, Schmidt PN, Gluud LL. Infection increases mortality in necrotizing pancreatitis: a systematic review and meta-analysis. Pancreatology. 2016;16(5):698–707.
3. Tandon RK. Management of acute pancreatitis: Indian guidelines and protocols. API Med Update. 2013;23:267–70.
4. Banks PA, Bollen TL, Dervenis C, Gooszen HG, Johnson CD, Sarr MG, et al. Classification of acute pancreatitis 2012: revision of the Atlanta classification and definitions by international consensus. Gut. 2013;62(1):102–11.
5. Thoeni RF. The revised Atlanta classification of acute pancreatitis: its importance for the radiologist and its effect on treatment. Radiology. 2012;262(3):751–64.
6. Walkowska J, Zielinska N, Tubbs RS, Podgórski M, Dłubek-Ruxer J, Olewnik Ł. Diagnosis and treatment of acute pancreatitis. Diagnostics (Basel). 2022;12(8):1974.
7. Siregar GA, Siregar GP. Management of severe acute pancreatitis. Open Access Maced J Med Sci. 2019;7(19):3319–23.
8. Ranson JH, Pasternack BS. Statistical methods for quantifying the severity of clinical acute pancreatitis. J Surg Res. 1977;22(2):79–91.
9. Theerthegowda AN, Umashankar P, Iyer NS. A comparative study between bedside index for severity in acute pancreatitis (BISAP) and acute physiology and chronic health evaluation (APACHE-II) scoring system in assessing the severity of acute pancreatitis. J Evid Based Med Healthc. 2021;8(36):3269–75.

10. Balthazar EJ, Robinson DL, Megibow AJ, Ranson JH. Acute pancreatitis: value of CT in establishing prognosis. Radiology. 1990;174(2):331–6.
11. Mortele KJ, Wiesner W, Intriere L, Shankar S, Zou KH, Kalantari BN, et al. A modified CT severity index for evaluating acute pancreatitis: improved correlation with patient outcome. AJR Am J Roentgenol. 2004;183(5):1261–5.
12. Wu BU, Johannes RS, Sun X, Tabak Y, Conwell DL, Banks PA. The early prediction of mortality in acute pancreatitis: a large population-based study. Gut. 2008;57(12):1698–703.
13. Singh VK, Wu BU, Bollen TL, Repas K, Maurer R, Johannes RS, et al. A prospective evaluation of the bedside index for severity in acute pancreatitis score in assessing mortality and intermediate markers of severity in acute pancreatitis. Am J Gastroenterol. 2009;104(4):966–71.
14. Khanna AK, Meher S, Prakash S, Tiwary SK, Singh U, Srivastava A, et al. Comparison of Ranson, Glasgow, MOSS, SIRS, BISAP, APACHE-II, CTSI scores, IL-6, CRP, and procalcitonin in predicting severity, organ failure, pancreatic necrosis, and mortality in acute pancreatitis. HPB Surg. 2013;2013:367581.
15. Chen L, Lu G, Zhou Q, Zhan Q. Evaluation of the BISAP score in predicting severity and prognoses of acute pancreatitis in Chinese patients. Int Surg. 2013;98(1):6–12.
16. Gupta AK, Raj S, Chaudhary P, Sharma A, Ranjan P. A prospective comparative study of bedside index for assessing severity in acute pancreatitis, APACHE II and computed tomography severity index scoring in predicting outcome in acute pancreatitis. Hellenic J Surg. 2015;87(6):473–8.
17. Jingoniya NK, Yadav BL, Verma PK, Bansal S, Gupta S. Comparative evaluation of BISAP score and computed tomography severity index as a predictor for severity of acute pancreatitis. Int Surg J. 2022;9(2):421–5.
18. Chatterjee R, Parab N, Sajjan B, Nagar VS. Comparison of Acute Physiology and Chronic Health Evaluation II, modified computed tomography severity index, and bedside index for severity in acute pancreatitis score in predicting the severity of acute pancreatitis. Indian J Crit Care Med. 2020;24(2):99–103.
19. Papachristou GI, Muddana V, Yadav D, O'Connell M, Sanders MK, Slivka A, et al. Comparison of BISAP, Ranson's, APACHE-II, and CTSI scores in predicting organ failure, complications, and mortality in acute pancreatitis. Am J Gastroenterol. 2010;105(2):435–41.
20. Cho JH, Kim TN, Chung HH, Kim KH. Comparison of scoring systems in predicting the severity of acute pancreatitis. World J Gastroenterol. 2015;21(8):2387–94.
21. Tahir H, Rahman S, Habib Z, Khan Y, Shehzad S. Comparison of the accuracy of modified CT severity index score and neutrophil-to-lymphocyte ratio in assessing the severity of acute pancreatitis. Cureus. 2021;13(8):e17020.